arXiv:0912.2436v1 [math.CO] 12 Dec 2009

On Gowers norms of some functions *

Shkredov I.D.

Annotation.

We consider a class of two–dimensional functions $f(x,y)$ with the property that the smallness of its rectangular norm implies the smallness of rectangular norm for $f(x,x+y)$. Also we study a family of functions $f(x,y)$ having a similar property for higher Gowers norms. The method based on a transference principle for a class of sums over special systems of linear equations.

1. Introduction.

The notion of Gowers norms was introduced in papers [2, 3] and is a very important tool of investigation in wide class of problems of additive combinatorics (see e.g. [2, 3] [8]—[20]) as well as in ergodic theory (see e.g. [15], [21]—[29]). Recall the definitions.

Let $\mathbf{G}$ be a finite set, and $N = |\mathbf{G}|$. Let also d be a positive integer, and $\{0,1\}^d = \{\omega = (\omega_1,\dots,\omega_d) \ : \ \omega_j \in \{0,1\}, j = 1,2,\dots,d\}$ be an ordinary d—dimensional cube. For $\omega \in \{0,1\}^d$ denote by $|\omega|$ the sum $\omega_1 + \dots + \omega_d$. Let also $\mathcal{C}$ be the operator of complex conjugation. Let $\vec{x} = (x_1,\dots,x_d), \vec{x}' = (x'_1,\dots,x'_d)$ be two arbitrary vectors from $\mathbf{G}^d$. By $\vec{x}^\omega = (\vec{x}^\omega_1,\dots,\vec{x}^\omega_d)$ denote the vector

$$\vec{x}^\omega_i = \begin{cases} x_i & \text{if } \omega_i = 0, \\ x'_i & \text{if } \omega_i = 1. \end{cases}$$

Thus $\vec{x}^\omega$ depends on $\vec{x}$ and $\vec{x}'$.

Let X be a non–empty finite set, $Z : X \to \mathbb{C}$ be a function. Denote by $\mathbb{E}Z = \mathbb{E}_x Z$ the sum $\frac{1}{|X|}\sum_{x\in X} Z(x)$.

Let $f : \mathbf{G}^d \to \mathbb{C}$ be an arbitrary function. We will write $f(\vec{x})$ for $f(x_1,\dots,x_d)$.

Definition 1.1 (see [2, 3]) Gowers U^d–norm (or d–uniformity norm) of the function f is the following expression

$$\|f\|_{U^d} = \left(\mathbb{E}_{\vec{x}\in\mathbf{G}^d}\, \mathbb{E}_{\vec{x}'\in\mathbf{G}^d} \prod_{\omega\in\{0,1\}^d} \mathcal{C}^{|\omega|} f(\vec{x}^\omega) \right)^{1/2^d} . \qquad (1)$$

*This work was supported Pierre Deligne's grant based on his 2004 Balzan prize, grant RFFI N 06-01-00383, grant of President of Russian Federation MK-1959.2009.1, and grant Leading Scientific Schools No. 691.2008.1.

Keywords : Gowers norms, linear equations.
MSC 2000 : 11B75, 11B99.

A sequence of 2^d points $\vec{x}^\omega$, $\omega \in \{0,1\}^d$ is called *d–dimensional cube*. Thus the summation in formula (1) is taken over all cubes of $\mathbf{G}^d$. For example, $\{(x,y),(x',y),(x,y'),(x',y')\}$, where $x,x',y,y' \in \mathbf{G}$ is a two–dimensional cube. In the case Gowers norm is called *rectangular norm*.

For $d=1$ the expression above gives a semi–norm but for $d \ge 2$ Gowers norm is a norm. In particular, the triangle inequality holds

$$\|f+g\|_{U^d} \le \|f\|_{U^d} + \|g\|_{U^d}\,. \tag{2}$$

One can prove also (see [3]) the following monotonicity relation. Let $f_{x_d}(x_1,\dots,x_{d-1}) := f(x_1,\dots,x_d)$. Then

$$\mathbb{E}_{x_d \in \mathbf{G}} \|f_{x_d}\|^{2^{d-1}}_{U^{d-1}} \le \|f\|^{2^{d-1}}_{U^d} \tag{3}$$

for all $d \ge 2$.

If $\mathbf{G} = (\mathbf{G},+)$ is a finite Abelian group with additive group operation $+$, $N = |\mathbf{G}|$ then one can "project" the norm above onto the group $\mathbf{G}$ and obtain the ordinary Gowers norm. In other words, we put the function $f(x_1,\dots,x_d)$ in formula (1) equals "one–dimensional" function $f(x_1,\dots,x_d) := f(x_1+\dots+x_d)$. Denoting the obtained norm as U^k and we have an analog of (3)

$$\|f\|_{U^{d-1}} \le \|f\|_{U^d} \tag{4}$$

for all $d \ge 2$.

Gowers norms have the following characteristic property. Let $d \ge 2$, $\vec{x} = (x_1,\dots,x_d) \in \mathbf{G}^d$ be an arbitrary vector and $i \in \{1,\dots,d\}$. By $(\vec{x})_{(i)}$ denote the vector $(\vec{x})_{(i)} = (x_1,\dots,x_{i-1},x_{i+1},\dots,x_d) \in \mathbf{G}^{d-1}$. Applying the Cauchy–Schwartz (see [3] and [4, 5]) several times, we have

Lemma 1.2 *Let $d \ge 2$ be an integer, and $f : \mathbf{G}^k \to \mathbb{C}$ be a function. Let also $u_1,\dots,u_d : \mathbf{G}^d \to [-1,1]$ be any functions such that $u_i(\vec{x}) = u_i((\vec{x})_{(i)})$, $i = 1,\dots,d$, $\vec{x} = (x_1,\dots,x_d)$. Then*

$$\left| \sum_{\vec{x}} f(\vec{x}) \prod_{i=1}^{d} u_i(\vec{x}) \right| \le \|f\|_{U^d}\,.$$

Thus any function with small Gowers norm does not correlate with product of any functions which depend on smaller number of variables.

In the paper we concentrate on the case of two–dimensional functions $f : \mathbf{G}\times\mathbf{G} \to \mathbb{C}$, $\mathbf{G}$ is a finite Abealian group. For any positive integer t one can consider U^{t+1}—norm of the function $F_t(x,y_1,\dots,y_t) := f(x, x+y_1+\dots+y_t)$ and the function $H_t(x,y_1,\dots,y_t) := f(x, y_1+\dots+y_t)$. It is easy to construct examples of functions $f(x,y)$ with, say, huge rectangular norm and small quantity $\|f(x,x+y)\|_{U^2}$ ("skew rectangular norm") and vice versa. Thus there is no obvious dependence between the numbers $\|F_t\|_{U^{t+1}}$ and $\|H_t\|_{U^{t+1}}$ in general. Nevertheless, we find a class of functions such that the smallness of $\|H_t\|_{U^{t+1}}$ implies the smallness $\|F_t\|_{U^{t+1}}$.

By $\widehat{\mathbf{G}}$ denote the Pontryagin dual of $\mathbf{G}$. In other words $\widehat{\mathbf{G}}$ is the group of homomorphisms ξ from $\mathbf{G}$ to $\mathbf{R}/\mathbf{Z}$, $\xi : x \to \xi \cdot x$. It is well–known that in the case of Abelian group $\mathbf{G}$ the dual group $\widehat{\mathbf{G}}$ is isomorphic to $\mathbf{G}$.

Let us formulate one of the main results of the paper.

Theorem 1.3 *Let $a : \mathbf{G} \to \mathbf{G}$ be a function and $f(x,y) = e(x \cdot a(y))$, where $e(x) = e^{2\pi i x}$. Then the condition*

$$\|f(x, y_1+\dots+y_t)\|_{U^{t+1}} = o(1)\,, \quad N \to \infty \tag{5}$$

implies

$$\|f(x, x+y_1+\dots+y_t)\|_{U^{t+1}} = o(1)\,, \quad N \to \infty \tag{6}$$

for all positive integers t.

More precisely, we obtain a quantitative form of the formulas above (see Corollary 3.3). Note that more strong condition than (5), namely, $\|f(x, y_1+\dots+y_{t+1})\|_{U^{t+1}} = o(1)$ trivially implies (6) (see proof of Corollary 3.3). Theorem 1.3 will be derived from the more general Theorem 2.4 of section 2. The result says, roughly speaking, that a class of sums of general form is taken over systems of linear equations including an arbitrary function $a(x)$ can be reduced to sums having no these linear restrictions (see Theorem 2.4). It is interesting we do not use Fourier analysis in the proof. Our main tool is a suitable version of Lemma 9.3 from Gowers' paper [3].

Let us say a few words about the notation. If $S \subseteq \mathbf{G}$ is a set then we will write $S(x)$ for the characteristic function. In other words $S(x) = 1$ if $x \in S$ and zero otherwise. By log denote logarithm base two and by $\mathbb{D}$ denote the unit disk on the complex plane. Sings $\ll$ and $\gg$ are usual Vinogradov's symbols. If n is a positive integer then we will write $[n]$ for the segment $\{1, 2, \dots, n\}$.

The author is grateful to N.G. Moshchevitin for useful discussions.

2. The proof of the main result.

We need in a number of definitions.

Let $l \geq 2$, m be any positive integers, $m < l$. Let also $x_1, \dots, x_l$, $y_1, \dots, y_l$ be some variables. Consider a system of linear equations

$$\sum_{j=1}^{l} \varepsilon_j^{(i)} x_j = 0\,, \quad i = 1, 2, \dots, m\,, \tag{7}$$

and also an equation

$$\sum_{j=1}^{l} \varepsilon_j^{(1)} y_j = 0\,. \tag{8}$$

Here $\varepsilon_j^{(i)} \in \{0, -1, 1\}$. So, if we know system (7) then we automatically know equation (8). Suppose that the rank of subsystem (7) equals m.

Let S be the family of all systems from m equations of full rank having the form $\sum_{j=1}^{l} \eta_j x_j = 0$, $\eta_j \in \{0, \pm 1\}$ such that η can be written as a linear combination of vectors $(\varepsilon_1^{(1)}, \dots, \varepsilon_l^{(1)}), \dots, (\varepsilon_1^{(m)}, \dots, \varepsilon_l^{(m)})$ from (7). For example, we can multiply some equations of (7) by -1 and get a system from S. Thus $|S| \geq 2^m$. On the other hand, $|S| \leq 3^{lm}$. Clearly, if a tuple $(x_1, \dots, x_l)$ is a solution of some system from S then it is a solution of any system from S, in particular, system (7). By Υ denote system (7), (8).

Let E be a family of linear equations with coefficient $\{0, \pm 1\}$ which can be written as a linear combinations of some equations from S. Clearly, $|E| \leq 3^l$. Suppose that $e \in E$ and define $\theta(e) = 2^t$, where t is the number of zero coefficients in e. Put also $\theta(v) = \prod_{e \in v} \theta(e)$, v is a system. Lastly, let

$$\theta = \theta_{m-1}(E) = \sum_{v} \theta(v)\,,$$

where summation is taken over all systems with $m-1$ equations from E.

Let $B_1, \dots, B_l \subseteq \mathbf{G}$ be arbitrary sets. Suppose that there are some maps $\varphi_1, \dots, \varphi_l$, $\varphi_j : B_j \to \mathbf{G}$, $j \in [l]$.

Finally, let C be a subset of $\mathbf{G}^l$. We will call a tuple $(x_1, \dots, x_l)$ satisfying a system from S *additive*. An additive tuple $(x_1, \dots, x_l)$ is called *degenerate* if there is a non–zero vector $\vec{\eta} = (\eta_1, \dots, \eta_l)$, $\eta_j \in \{0, \pm 1\}$ such that $\sum_{j=1}^l \eta_j x_j = 0$ and the last equation does not belong E. Otherwise, such a tuple will be called *non-degenerate*. Note that non–degenerate additive tuples can satisfy equations from E only. Further, non-degenerate additive tuple $(x_1, \dots, x_l) \in C$ is called *good* if $(x_1, \dots, x_l)$, $(\varphi_1(x_1), \dots, \varphi_l(x_l))$ satisfies (7), (8) for some system from S, and *bad* otherwise. So, by definition, good and bad tuples are non–degenerate and belong to C.

First of all we prove the following simple extension of Lemma 9.3 from [3].

Lemma 2.1 *Let $\alpha, \omega, \eta \in (0, 1]$ be any numbers, $B_j \subseteq \mathbf{G}$ be arbitrary sets, and C be a subset of $\mathbf{G}^l$. Suppose that T is a parameter, the number of additive tuples in $(B_1 \times \ldots \times B_l) \bigcap C$ is at most $\omega^{-1} T$,*

$$T \ge 2^{2l} |S| \left(\frac{2(1+\omega^{-1})}{\alpha\eta} \right)^{lm2^{3ml}} N^{l-m-1} \,, \tag{9}$$

and there exist at least αT good tuples. Then there are some sets $B'_j \subseteq B_j$ such that the number of good tuples $(x_1, \dots, x_l) \in (B'_1 \times \ldots \times B'_l) \bigcap C$ is at least

$$\left(\frac{\alpha\eta}{2(1+\omega^{-1})} \right)^{lm2^{3ml}} T$$

and the ratio of good tuples in $(B'_1 \times \ldots \times B'_l) \bigcap C$ to the number of bad tuples in $(B'_1 \times \ldots \times B'_l) \bigcap C$ is at least $(1-\eta)$.

Proof. Let k be a natural parameter and choose $r_1, \dots, r_k$, $s_1, \dots, s_k$, $w^{(1)}_1, \dots, w^{(1)}_k, \dots, w^{(m-1)}_1, \dots, w^{(m-1)}_k \in \mathbf{G}$ uniformly and independently. After that independently choose points $x_j \in B_j$ such that a point x_j go into a new set B'_j with probability

$$p(x_j) = \frac{1}{2^{mk}} \prod_{i=1}^k (1 + \mathrm{cs}\,(r_i x_j + s_i \varphi_j(x_j))) \cdot \prod_{q=1}^{m-1} \left(1 + \mathrm{cs}\,(w^{(q)}_i x_j) \right) \,,$$

where $\mathrm{cs}\,(x) = \cos(\frac{2\pi}{N} x)$. A tuple $(x_1, \dots, x_l)$ belongs to $B'_1 \times \ldots \times B'_l$ with probability

$$\frac{1}{2^{mlk} N^{(m+1)k}} \sum_{r_1, \dots, r_k} \sum_{s_1, \dots, s_k} \sum_{w^{(1)}_1, \dots, w^{(1)}_k, \dots, w^{(m-1)}_1, \dots, w^{(m-1)}_k} \prod_{i=1}^k \prod_{j=1}^l (1 + \mathrm{cs}\,(r_i x_j + s_i \varphi_j(x_j))) \times$$

$$\times \prod_{q=1}^{m-1} \left(1 + \mathrm{cs}\,(w^{(q)}_i x_j) \right) =$$

$$= \frac{1}{2^{mlk} N^{(m+1)k}} \left(\sum_{r,s} \sum_{w_1, \dots, w_{m-1}} \prod_{j=1}^l (1 + \mathrm{cs}\,(r x_j + s \varphi_j(x_j))) \prod_{q=1}^{m-1} (1 + \mathrm{cs}\,(w_q x_j)) \right)^k =$$

$$= \frac{1}{2^{mlk} N^{(m+1)k}} (2^{-lm} \sum_{r,s} \sum_{w_1, \dots, w_{m-1}}$$

$$\prod_{j=1}^{l} (2 + e(rx_j + s\varphi_j(x_j)) + e(-rx_j - s\varphi_j(x_j))) \prod_{q=1}^{m-1} (2 + e(w_q x_j) + e(-w_q x_j)))^k =$$

$$= \frac{1}{2^{mlk} N^{(m+1)k}} (2^{-lm} \sum_{r,s} \sum_{w_1,\dots,w_{m-1}} \prod_{j=1}^{l} \sum_{\varepsilon_1^{(1)},\dots,\varepsilon_l^{(1)},\dots,\varepsilon_1^{(m)},\dots,\varepsilon_l^{(m)} \in \{0,\pm 1\}} 2^{|\{i,j \,:\, \varepsilon_j^{(i)}=0\}|} \times$$

$$\times e(r \sum_j \varepsilon_j^{(1)} x_j + s \sum_j \varepsilon_j^{(1)} \varphi_j(x_j) + w_1 \sum_j \varepsilon_j^{(2)} x_j + \dots + w_{m-1} \sum_j \varepsilon_j^{(m)} x_j))^k \,.$$

Thus, the last probability does not equal zero if there is a tuple $(\varepsilon_j^{(i)})$, $\varepsilon_j^{(i)} \in \{0, \pm 1\}$ such that

$$\sum_j \varepsilon_j^{(1)} x_j = \sum_j \varepsilon_j^{(1)} \varphi_j(x_j) = \sum_j \varepsilon_j^{(2)} x_j = \dots = \sum_j \varepsilon_j^{(m)} x_j = 0 \,. \tag{10}$$

It is so, for example, if all $\varepsilon_j^{(i)}$ equal zero.

Let $(x_1, \dots, x_l)$ be a bad tuple. Then, by definition, $(x_1, \dots, x_l)$ satisfy a system $\upsilon \in S$ and is non–degenerate. It is easy to see that there is no vector $(\varepsilon_1^{(1)}, \dots, \varepsilon_l^{(1)}) \neq \vec{0}$ such that (10) holds. Indeed, otherwise we can add some equations from υ and obtain a contradiction with the fact that $(x_1, \dots, x_l)$ is a bad sequence. Thus, every bad additive tuple is chosen with probability $2^{-lmk}(2^{-lm} 2^l \theta)^k$. Clearly, $\theta \geq 2^{l(m-1)}$ and put $\theta = 2^{l(m-1)} + \theta_1$, $\theta_1 \geq 0$. Then the last probability equals $2^{-lmk}(1 + \theta_1 2^{-l(m-1)})^k$.

On the other hand, it is easy to see that a good additive tuple is chosen with probability at least

$$2^{-lmk}(2^{-lm}(2^l \theta + 2\theta))^k \geq 2^{-lmk}(1 + \theta_1 2^{-l(m-1)} + 2^{-(l-1)})^k > 2^{-lmk}(1 + \theta_1 2^{-l(m-1)})^k \,.$$

Note that there are at most $3^l |S| |B_1| \dots |B_{l-m-1}| \leq 3^l |S| N^{l-m-1}$ degenerate tuples. Now, let X and Y be the numbers of good and bad additive tuples in $B_1' \times \dots \times B_l'$. Using the assumption of the lemma, we get

$$\mathbb{E}X \geq 2^{-lmk}(1 + \theta_1 2^{-l(m-1)} + 2^{-(l-1)})^k \alpha T$$

and

$$\mathbb{E}Y \leq 2^{-lmk}(1 + \theta_1 2^{-l(m-1)})^k \omega^{-1} T \,.$$

It is easy to see that $\theta_1 < \theta \leq 3^{lm} 2^{lm}$. Since $(1 + 2^{-3lm})^k \geq 2^{k 2^{-3lm}}$ it follows that

$$\left(\frac{1 + \theta_1 2^{-l(m-1)} + 2^{-(l-1)}}{1 + \theta_1 2^{-l(m-1)}} \right)^k > (1 + 2^{-3lm})^k \geq 2^{k 2^{-3lm}}$$

and the last expression is at least $(1 + \omega^{-1})/(\alpha\eta)$, provided by we choose an integer k from the conditions

$$2 \left(\frac{\alpha\eta}{2(1 + \omega^{-1})} \right)^{2^{3lm}} \leq 2^{-k} \leq \left(\frac{\alpha\eta}{1 + \omega^{-1}} \right)^{2^{3lm}} \,. \tag{11}$$

Hence

$$\eta \mathbb{E}X - \mathbb{E}Y \geq \alpha\eta 2^{-lmk}(1 + \theta_1 2^{-l(m-1)} + 2^{-(l-1)})^k T - 2^{-lmk}(1 + \theta_1 2^{-l(m-1)})^k \omega^{-1} T \geq 2^{-lmk} T \,. \tag{12}$$

Using (9), (11), we have

$$2^{-lmk}T \geq 2\left(\frac{\alpha\eta}{2(1+\omega^{-1})}\right)^{lm2^{3ml}} T \geq 2\cdot 4^l|S|N^{l-m-1} \geq 2(3^l|S||B_1|\dots|B_{l-m-1}|)\,. \qquad (13)$$

By (12), (13), we obtain that there are sets $B_1' \subseteq B_1, \dots, B_l' \subseteq B_l$ such that $\eta X \geq Y$ and $X \geq \left(\frac{\alpha\eta}{2(1+\omega^{-1})}\right)^{lm2^{3ml}} T$.This completes the proof.

Note 2.2 Certainly, one can suppose that our tuples $(x_1,\dots,x_l)$, $(\varphi(x_1),\dots,\varphi(x_l))$, $x_j \in B_j'$ satisfying (7) and (8) for some system υ from S, satisfy a new system

$$\sum_{j=1}^{l} \bar{\varepsilon}_j^{(i)} x_j = 0\,, \quad i = 1,2,\dots,m\,, \qquad (14)$$

and

$$\sum_{j=1}^{l} \bar{\varepsilon}_j^{(1)} \varphi(x_j) = 0\,, \qquad (15)$$

where $\bar{\varepsilon}_j^{(i)} \in \{0,\pm 1\}$, $\bar{\varepsilon}_j^{(i)} = \varepsilon_j^{(i)}$ for any $j \geq 1$, $i \geq 2$, the rank of system (14) equals m, and $\bar{\varepsilon}_j^{(1)}$ depend on υ.

Definition 2.3 Let L, l be positive integers. A set $\Omega \subseteq \mathbf{G}^l$ is called a *set of level* L if

$$\Omega = \prod_{j=1}^{L} (\Omega_j' - \Omega_j'')$$

where each Ω_j', Ω_j'', $\Omega_j'' \subseteq \Omega_j'$ is a Cartesian product of some sets. We say that Cartesian products are sets of level zero.

Let C be a set, and let f be a complex function which depends on variables $(x_1,\dots,x_l) \in (B_1 \times \dots \times B_l) \cap C$, $\vec{x} = (x_1,\dots,x_l)$ and f also depends on $\varphi_1(x_1),\dots,\varphi_l(x_l)$, $\varphi(\vec{x}) = (\varphi_1(x_1),\dots,\varphi_l(x_l))$. Let υ be a system from S and

$$\sigma_{f,\upsilon}(B_1,\dots,B_l;C) := \sum_{(\vec{x},\varphi(\vec{x}))\ \text{satisfies}\ (7),(8)\ \text{for}\ \upsilon\in S} f(x_1,\dots,x_l,\varphi_1(x_1),\dots,\varphi_l(x_l))\,. \qquad (16)$$

Further $\sigma_f(B_1,\dots,B_l;C) = \sum_{\upsilon\in S} |\sigma_{f,\upsilon}(B_1,\dots,B_l;C)|$. Take an arbitrary number $\rho \in [l]$, express x_ρ from (7), $\varphi_\rho(x_\rho)$ from (8) and substitute x_ρ, $\varphi_\rho(x_\rho)$ into f. After that choose another $m-1$ variables $x_{j_1},\dots,x_{j_{m-1}}$ express them from (7) and substitute into f. Let $\vec{j} = (j_1,\dots,j_{m-1})$. We get a new function

$$f^{\upsilon}_{\rho,\vec{j}}(\vec{x}) := f^{\upsilon}_{\rho,\vec{j}}(x_j,\varphi_j(x_j))\,, \quad \rho \notin \{j_1,\dots,j_{m-1}\}\,.$$

Let $\|f\|_{\mathcal{U}^d} = N^{2d}\|f\|_{U^d}^{2^d}$ and for any $\rho_1,\rho_2 \in [d]$, $\rho_1 < \rho_2$, we put

$$\|f\|_{\mathcal{U}^2(\rho_1,\rho_2)} = \sum_{x_1,\dots,x_d} \sum_{x'_{\rho_1},x'_{\rho_2}} f(x_1,\dots,x_{\rho_1},\dots,x_{\rho_2},\dots,x_d)\overline{f(x_1,\dots,x'_{\rho_1},\dots,x_{\rho_2},\dots,x_d)}\times$$

$$\times\overline{f(x_1,\dots,x_{\rho_1},\dots,x'_{\rho_2},\dots,x_d)}f(x_1,\dots,x'_{\rho_1},\dots,x'_{\rho_2},\dots,x_d)\,.$$

Formulate the main result of the section.

Theorem 2.4 *Let $\varepsilon \in (0,1]$ be a real number, $\varepsilon \le 2^{-2^{20}l^2m^2 2^{6ml}}$, and f be a complex function, $f : \mathbf{G} \to \mathbb{D}$. Let also $B_1, \dots, B_l$ be arbitrary sets. Suppose that for any $\upsilon \in S$ there are $\vec{j}$, $\rho \in [l]$ and $\rho_1 \ne \rho, \rho_2 \ne \rho$, $\rho_1 \ne \rho_2$ such that*

$$\| f^{\upsilon}_{\rho,\vec{j}} (\vec{x}) \|_{\mathcal{U}^2(\rho_1,\rho_2)} \le \varepsilon N^{l-m+2} \,. \tag{17}$$

Then

$$\sigma_f (B_1,\dots,B_l; \mathbf{G}^l) \le \max \left\{ 2^{30} \left(\frac{512}{\log(1/\varepsilon)} \right)^{(128lm2^{3ml})^{-1}} , 4 \left(\frac{2^{2l}}{N} \right)^{(16lm2^{3ml})^{-1}} \right\} 3^{lm} N^{l-m} \,. \tag{18}$$

Proof. Let ε_1 be the maximum in the right hand side of (18) divided by N^{l-m}. Let also $\upsilon \in S$ and denote by $\tau_{f,\upsilon} (B_1,\dots,B_l; C)$ the number of solutions of system (14), (15). By $\tau_f (B_1,\dots,B_l; C)$ denote the sum $\sum_{\upsilon \in S} \tau_{f,\upsilon} (B_1,\dots,B_l; C)$. Suppose that inequality (18) does not hold. Since $\tau_f (B_1,\dots,B_l; \mathbf{G}^l) \le |S| N^{l-m}$ and

$$\sigma_f (B_1,\dots,B_l; \mathbf{G}^l) \le \tau_f (B_1,\dots,B_l; \mathbf{G}^l) \,, \tag{19}$$

it follows that $\sigma_f (B_1,\dots,B_l; \mathbf{G}^l) \ge \varepsilon_1 |S|^{-1} \tau_f (B_1,\dots,B_l; \mathbf{G}^l)$ and $\tau_f (B_1,\dots,B_l; \mathbf{G}^l) \ge \varepsilon_1 N^{l-m}$.

We need in a "density increment" lemma.

Lemma 2.5 *Let $\alpha, \omega \in (0,1]$ be any numbers, B_j be arbitrary sets, and C be a set of level L. Suppose that*

$$\tau_f (B_1,\dots,B_l; C) \ge |S| \max\{ 2^{2l} \left(\frac{32(1+\omega^{-1})|S|}{\alpha\varepsilon_1} \right)^{lm2^{3ml}} N^{l-m-1}, \omega N^{l-m} \} \,, \tag{20}$$

$$\varepsilon \le 2^{-4L-16} |S|^{-8} \varepsilon_1^4 \left(\frac{\alpha\varepsilon_1}{32(1+\omega^{-1})|S|} \right)^{4lm2^{3ml}} \left(\frac{\tau_f (B_1,\dots,B_l; C)}{N^{l-m}} \right)^4 \tag{21}$$

and $f : \mathbf{G} \to \mathbb{D}$ is a function satisfying (17). Let also

$$\sigma_f (B_1,\dots,B_l; C) \ge \alpha \tau_f (B_1,\dots,B_l; C) \,. \tag{22}$$

Then there is a set $\tilde{C} \subseteq C$ of level at most $L+1$ and there are sets B'_j, $j \in [l]$, $(B'_1 \times \dots \times B'_l) \bigcap \tilde{C} = \emptyset$ such that

$$\sigma_f (B_1,\dots,B_l; C) \le \sigma_f (B_1,\dots,B_l; \tilde{C}) + 2^{-2} \varepsilon_1 |S|^{-1} \tau_f (B'_1,\dots,B'_l; C) \,. \tag{23}$$

and

$$\tau_f (B_1,\dots,B_l; \tilde{C}) \le (1-\zeta) \tau_f (B_1,\dots,B_l; C) \,, \tag{24}$$

where $\zeta = \left(\frac{\alpha\varepsilon_1}{32(1+\omega^{-1})|S|} \right)^{lm2^{3ml}}$.

Proof. Using Lemma 2.1 with parameters $\alpha, \omega, \eta = |S|^{-1}\varepsilon_1/16$ and $T = \tau_f (B_1,\dots,B_l; C)$, we get the sets $B'_j \subseteq B_j$, $j \in [l]$ such that the number g of good tuples $(x_1,\dots,x_l) \in (B'_1 \times \dots \times B'_l) \bigcap C$ is at least

$$g \ge \left(\frac{\alpha\eta}{2(1+\omega^{-1})} \right)^{lm2^{3ml}} \tau_f (B_1,\dots,B_l; C) = \zeta \tau_f (B_1,\dots,B_l; C) \,. \tag{25}$$

Thus
$$\tau_f(B'_1,\ldots,B'_l;C) \ge g \ge \zeta\tau_f(B_1,\ldots,B_l;C)\,. \tag{26}$$
Let $(B'_j)^1 = B'_j$ and $(B'_j)^0 = B_j \setminus B'_j$. Then for any $\upsilon \in S$, we have
$$\sigma_{f,\upsilon}(B_1,\ldots,B_l;C) = \sum_{\varpi\in\{0,1\}^l} \sigma_{f,\upsilon}((B'_1)^{\varpi_1},\ldots,(B'_l)^{\varpi_l};C) = \sum_{\varpi\in\{0,1\}^l} \sigma_{f,\upsilon}(\varpi;C)\,,$$
where $\varpi = (\varpi_1,\ldots,\varpi_l)$. Without losing of generality, assume that $\rho = l$ and $\vec{j} = (x_{l-m+1},\ldots,x_{l-1})$. Consider the term $\sigma_1(\upsilon) := \sigma_{f,\upsilon}(B'_1,\ldots,B'_l;C)$ which corresponds to $\varpi_1 = \cdots = \varpi_l = 1$. We have
$$\sum_{\upsilon\in S} |\sigma_1(\upsilon)| \le \sum_{\upsilon\in S} \left| \sum_{\vec{x}\in(B'_1\times\ldots\times B'_l)\bigcap C,\,(\vec{x},\varphi(\vec{x}))\text{ satisfies (7),(8) for }\upsilon} f^{\upsilon}_{\rho,\vec{j}}(\vec{x}) \right| +$$
$$+\eta\tau_f(B'_1,\ldots,B'_l;C) = \sigma'_1 + \sigma''_1\,. \tag{27}$$
It is easy to see, using (17) and Lemma 1.2 that
$$\sigma'_1 \le 2^L|S|\varepsilon^{1/4}N^{l-m} \le 2^L 3^{lm}\varepsilon^{1/4}N^{l-m}\,. \tag{28}$$
Put $\tilde{C} = ((B_1\times\ldots\times B_l)\setminus(B'_1\times\ldots\times B'_l))\bigcap C$. Clearly, $\tilde{C}$ is a set of level $L+1$. Using inequalities (27), (28), we obtain
$$\sigma_f(B_1,\ldots,B_l;C) = \sum_{\upsilon\in S} |\sigma_{f,\upsilon}(B_1,\ldots,B_l;C)| \le \sigma'_1 + \sigma''_1 + \sigma_f(B_1,\ldots,B_l;\tilde{C}) \le$$
$$\le 2^L 3^{lm}\varepsilon^{1/4}N^{l-m} + \eta\tau_f(B'_1,\ldots,B'_l;C) + \sigma_f(B_1,\ldots,B_l;\tilde{C})\,. \tag{29}$$
Recalling that $\eta = |S|^{-1}\varepsilon_1/16$, using condition (21) and formulas (26), (29), we get
$$\sigma_f(B_1,\ldots,B_l;C) \le \sigma_f(B_1,\ldots,B_l;\tilde{C}) + 2^{-2}\varepsilon_1|S|^{-1}\tau_f(B'_1,\ldots,B'_l;C)\,. \tag{30}$$
Finally,
$$\tau_f(B_1,\ldots,B_l;\tilde{C}) + \tau_f(B'_1,\ldots,B'_l;C) = \tau_f(B_1,\ldots,B_l;C)$$
and by (25), we obtain (24). This concludes the proof of the lemma.

Now return to the proof of the theorem. We use Lemma 2.5 inductively. At zeroth step, we have $\alpha \ge |S|^{-1}\varepsilon_1$ and ω is any positive number such that $\omega \le \varepsilon_1|S|^{-1}$. Suppose that our algorithm was applied h times. Thus we obtain the sets C_j, $j\in[h]$, every C_j is a set of level j and two sequences of $\zeta_1,\ldots,\zeta_h$, $\omega_1,\ldots,\omega_h$, each ζ_j depends on ω_j. Using inequality (23), we see that for any j the following holds
$$\sigma_f(B_1,\ldots,B_l;\mathbf{G}^l) \le \sigma_f(B_1,\ldots,B_l;C_j) + 2^{-2}\varepsilon_1|S|^{-1}\tau_f(B_1,\ldots,B_l;\mathbf{G}^l) \le$$
$$\le \sigma_f(B_1,\ldots,B_l;C_j) + 2^{-2}\varepsilon_1 N^{l-m}\,.$$
So, we can assume that for all $j\in[h]$, we have $\sigma_f(B_1,\ldots,B_l;C_j) \ge \varepsilon_1(2|S|)^{-1}\tau_f(B_1,\ldots,B_l;C_j)$. Further, in view of (19), we can suppose that for all $j\in[h]$ the following holds $\tau_f(B_1,\ldots,B_l;C_j) \ge 2^{-1}\varepsilon_1 N^{l-m}$. Thus for any $j\in[h]$, we can take

$\omega_j := 2^{-1}\varepsilon_1|S|^{-1}$ and hence for all $j \in [h]$ the following holds $\zeta_j \ge \left(\frac{\varepsilon_1^2}{64(1+2\varepsilon_1^{-1}|S|)|S|^2}\right)^{lm2^{3ml}} = \zeta_*$. By (24), we get

$$\tau_f(B_1,\dots,B_l;C_j) \le (1-\zeta_*)^j \tau_f(B_1,\dots,B_l;\mathbf{G}^l) \le (1-\zeta_*)^j|S|N^{l-m}$$

and our algorithm must stop after at most $2^4\log(2|S|\varepsilon_1^{-1})\left(\frac{\varepsilon_1^2}{64(1+2\varepsilon_1^{-1}|S|)|S|^2}\right)^{-lm2^{3ml}} := h$ number of steps. Clearly, the maximal level of any set which appears in our induction procedure does not exceed h. Using the inequality $\tau_f(B_1,\dots,B_l;C_h) \ge 2^{-1}\varepsilon_1 N^{l-m}$, we get from (21) the dependence between parameters ε and ε_1

$$\varepsilon \le 2^{-4h-20}|S|^{-8}\varepsilon_1^8\left(\frac{\varepsilon_1^2}{64(1+2\varepsilon_1^{-1}|S|)|S|^2}\right)^{4lm2^{3ml}} \tag{31}$$

Besides, by (20), we obtain

$$N \ge 2^{2l+1}\left(\frac{64(1+2\varepsilon_1^{-1}|S|)|S|^2}{\varepsilon_1^2}\right)^{lm2^{3ml}}|S|\varepsilon_1^{-1}. \tag{32}$$

If (31) and (32) hold then we have a contradiction. Small calculation shows that the last two inequalities imply (18). This completes the proof of Theorem 2.4.

Note 2.6 Probably, the result above suggests that we have a phenomenon in spirit of the well–known sum–product phenomenon (see e.g. [6] or [7]) or the dichotomy phenomenon (see e.g. [14]). Indeed, Theorem 2.4 is trivial in two opposite cases : if $a(x)$ is a linear function then $\sigma_f(B_1,\dots,B_l;\mathbf{G})$ is small by (17) and if $a(x)$ is far from all linear functions (e.g. $a(x)=x^2$) then (18) is small by (19) and an appropriate upper bound for $\tau_f(B_1,\dots,B_l;\mathbf{G})$. So, in some sense our result can be interpreted that there is no function linear and non–linear simultaneously.

3. Applications.

First of all let us note a simple property of Gowers norms (see e.g. [3]).

Lemma 3.1 *Let $d \ge 2$ be an integer, and $f : \mathbf{G}^k \to \mathbb{C}$ be a function. Let also $u_1,\dots,u_d : \mathbf{G}^d \to [-1,1]$ be any functions such that $u_i(\vec{x}) = u_i((\vec{x})_{(i)})$, $i=1,\dots,d$, $\vec{x}=(x_1,\dots,x_d)$. Then*

$$\|f(\vec{x})e(\prod_{i=1}^d u_i(\vec{x}))\|_{U^d} = \|f(\vec{x})\|_{U^d}.$$

Also, we need in a convexity lemma.

Lemma 3.2 *Let $\kappa > 0$ and $h(x) = 1/(\log x)^\kappa$. Then for any real numbers $x_1,\dots,x_N > 1$, we have*

$$\frac{1}{N}\sum_{i=1}^N h(x_i) \le h\left(\frac{1}{N}\sum_{i=1}^N x_i\right).$$

Now we obtain a corollary of Theorem 2.4. Clearly, the result below implies Theorem 1.3.

Corollary 3.3 *Let $\varepsilon \in (0,1)$ be a real number, $a : \mathbf{G} \to \mathbf{G}$ be a function, and $f(x,y) = e(x\cdot a(y))$. Suppose that*

$$\|f(x,y)\|_{U^2} \le \varepsilon. \tag{33}$$

Then

$$\|f(x,x+y)\|_{\mathcal{U}^2} \le 3^4 \max\left\{ 2^{30}\left(\frac{2^8}{\log(1/\varepsilon)}\right)^{2^{-21}}, 4\left(\frac{2^8}{N}\right)^{2^{-18}} \right\} N^4\,. \tag{34}$$

If t is a positive integer and

$$\|f(x,y_1+\dots+y_t)\|_{U^{t+1}} \le \varepsilon \tag{35}$$

then

$$\|f(x,x+y_1+\dots+y_t)\|_{\mathcal{U}^{t+1}} \ll_t \max\left\{ \frac{1}{\log(1/\varepsilon)}, N^{-c(t)} \right\} N^{2t+2}\,, \tag{36}$$

where $c(t)>0$ some constant depends on t only.

Note 3.4 Clearly, formula (34) of the corollary above holds for any function $f_\lambda(x,y) = e((x-\lambda)\cdot a(y))$, where $\lambda \in \mathbf{G}$ is an arbitrary element.

Proof. For $w=(w_1,\dots,w_d)$, $d\in\mathbb{N}$ and $\omega\in\{0,1\}^d$ we write w_ω for $\sum_{i=1}^d w_i^{\omega}$. Any sequence of points $(w_\omega)_{\omega\in\{0,1\}^d} \in \mathbf{G}$ is called *d–dimensional cube* (see [3]). So, we numerate the points of an arbitrary d–dimensional cube by index $\omega\in\{0,1\}^d$.

First of all let us note that the dimension of all d—dimensional cubes equals $d+1$ and any d—dimensional cube $(z_1,\dots,z_{2^d})$ is a solution of a system of full rank having $\sum_{i=0}^{d-2}\binom{d}{i} = 2^d-d-1$ linear equations, say

$$\begin{cases} \sum_{\omega\in\{0,1\}^d} z_\omega\cdot(-1)^{|\omega|} = 0\,, \\ \sum_{\omega\in\{0,1\}^d} z_\omega\cdot(-1)^{|\omega|} = 0\,, \quad \omega_i = 0,\ i\in[d]\,. \\ \dots \\ \sum_{\omega\in\{0,1\}^d} z_\omega\cdot(-1)^{|\omega|} = 0\,, \quad \omega_{i_1} = \dots = \omega_{i_{d-2}} = 0,\ i_j\in[d]\,. \end{cases}$$

By S denote the last system. The fact that S has rank equals 2^d-d-1 can be obtained by successfully considering of its equations from the first to the last (see variables $z_{1,\dots,1}$, $z_{0,1,\dots,1},\dots,z_{1,\dots,1,0}$ and so on).

Our task is count the quantity

$$\sigma_t := \|f(x,x+y_1+\dots+y_t)\|_{\mathcal{U}^{t+1}} = \sum_{x,x'}\sum_{y_1,\dots,y_t}\sum_{y'_1,\dots,y'_t} e(\sum_{\omega\in\{0,1\}^{t+1}} x_{\omega_1}\cdot a(x_{\omega_1}+y_{(\omega_2,\dots,\omega_{t+1})}))\,,$$

where $y=(y_1,\dots,y_t)$. Let us change the variables $x\to x-y_1$, $x'\to x'-y_1$, $y'_1-y_1=\Delta$. Let $z=(x,\Delta,y_2,\dots,y_t)$, $\omega\in\{0,1\}^{t+1}$, $\omega=(\bar\omega_1,\eta,\bar\omega_2)$, where $\bar\omega_1,\eta\in\{0,1\}$, $\bar\omega_2\in\{0,1\}^{t-1}$. Put $z_\omega = x_{\bar\omega_1}+\eta\Delta+y_{\bar\omega_2}$. So, z_ω depends on x,Δ,y. We write $\eta=\eta(z_\omega)$. Clearly,

$$\sigma_t = \sum_{x,x',\Delta}\ \sum_{y_2,\dots,y_t,\,y'_2,\dots,y'_t} e(\sum_{\omega\in\{0,1\}^{t+1}} x_{\bar\omega_1}\cdot a(x_{\bar\omega_1}+\eta\Delta+y_{(\omega_2,\dots,\omega_t)}))\times$$

$$\times N\delta_0(\sum_{\omega\in\{0,1\}^{t+1}} (-1)^{|\omega|}a(z_\omega))\,, \tag{37}$$

where $\delta_0(x)=1$ if $x=0$ and zero otherwise. It is easy to see that the sequence $(z_\omega)_{\omega\in\{0,1\}^{t+1}}$ form a cube. So, our sum σ_t has form (16). To prove Corollary 3.3 we need to check condition (17) but firstly let us prove the corollary in the simplest case $t=1$. In the situation formula (37) is

$$\sigma_1 = \sum_{x,x',\Delta} e(xa(x)-x'a(x')-xa(x+\Delta)+x'a(x'+\Delta))\cdot N\delta_0(a(x)-a(x')-a(x+\Delta)+a(x'+\Delta))$$

$$= \sum_{x,x',\Delta} e((x-x')(a(x)-a(x+\Delta))) \cdot N\delta_0(a(x)-a(x')-a(x+\Delta)+a(x'+\Delta))$$

and condition (17) holds because (33) and S contains just two equations in the case. Thus, we prove (34), provided by $t=1$.

Now suppose that $t \ge 2$. Let $\sum_\omega a(z_\omega)\bar{\eta}_\omega = 0$ be an equation from S. Suppose that there is $\bar{\eta}_\omega \ne 0$ such that $\eta(z_\omega) = 1$ (see definition of z_ω). Without losing of generality, one can suppose that $\bar{\eta}_{\vec{1}} = -1$, where $\vec{1} = (1,\dots,1)$. Then $a(z_{\vec{1}}) = \sum_{\omega,\, \omega \ne \vec{1}} a(z_\omega)\bar{\eta}_\omega$ and we can substitute it into formula (37). After that change the variables $x \to x - y_2$, $x' \to x' - y_2$, $y_2' - y_2 = q$. Consider the variables $x', q, y_3', \dots, y_t'$. We see there are terms $e(x'a(x+q+y_3'+\dots+y_t'))$, $e(x'a(x+\Delta+q+y_3'+\dots+y_t'))$ and may be the term $e((x'-y_2)a(x'+q+y_3'+\dots+y_t'))$, containing $x', q, y_3', \dots, y_t'$ but there is no term $e((x'-y_2)a(x'+\Delta+q+y_3'+\dots+y_t'))$ in the expression for σ_t. Put

$$\Theta^{(1)}_{y_2,x,\Delta}(x', q, y_3', \dots, y_t') = e(x'a(x+q+y_3'+\dots+y_t') + x'a(x+\Delta+q+y_3'+\dots+y_t')) \quad (38)$$

and

$$\Theta^{(2)}_{y_2,x,\Delta}(x', q, y_3', \dots, y_t') = \Theta^{(1)}_{y_2,x,\Delta}(x', q, y_3', \dots, y_t') \cdot e((x'-y_2)a(x'+q+y_3'+\dots+y_t')) \quad (39)$$

provided by the term $e((x'-y_2)a(x'+q+y_3'+\dots+y_t'))$ exists. Clearly, another multiples do not contain these variables. Using Lemma 3.1 and summation over y_2, we get

$$\sum_{y_2,x,\Delta} \|\Theta^{(2)}_{y_2,x}(x', q, y_3', \dots, y_t')\|_{\mathcal{U}^t} \le N^2 \|f(x, y_1+\dots+y_t)\|_{\mathcal{U}^{t+1}}\,.$$

Similarly, summing over x, Δ, we have

$$\sum_{y_2,x,\Delta} \|\Theta^{(1)}_{y_2,x}(x', q, y_3', \dots, y_t')\|_{\mathcal{U}^t} \le N^2 \|f(x, y_1+\dots+y_t)\|_{\mathcal{U}^{t+1}}\,.$$

By Lemma 3.2, we obtain (36). Now suppose that for our equation $\sum_\omega a(z_\omega)\bar{\eta}_\omega = 0$ from S for any $\bar{\eta}_\omega \ne 0$ we have $\eta(z_\omega) = 0$. But changing the variables $y_2 \to y_2 - \Delta$, $y_2' \to y_2' - \Delta$ we reducing our expression to the previous case. This concludes the proof.

Note 3.5 If we take $a(x) = \mathrm{const}$ then, clearly, (34) does not hold. On the other hand it is easy to see that (33) is also has no place. If we put $a(x) = x$ then (36) and (35) do not hold for $t = 2$. For $t \ge 2$ take $a(x) = x^{t-1}$. Then the norm of the functions from (35) and (36) equal one. Thus, in some sense our conditions (33), (35) are necessary.

Note 3.6 Let us analyze a little bit more general functions of the form $f(z_1+\dots+z_r+x_1+\dots+x_s, x_1+\dots+x_s+y_1+\dots+y_t)$, r is a positive integer, $s,t \ge 0$. If $r \ge 2$ then $\|f\|_{U^{s+t+r}} = 1$ for any function $a(x)$. Further, if $r=0$, $t=0$ and $\mathbf{G} = \mathbb{Z}_p$, p is a prime number then put $a(x) = x^{-1}$. It is easy to see that inequality (33) takes place (e.g. see Corollary 3.7) and $\|f\|_{U^s} = 1$. The case $r = s = 0$ is not interesting and consider the situation $r = 0$, $s \ge 2$, $t \ge 1$. It is easy to see that $\|e(xa(y_1+\dots+y_{s+t})\|_{U^{s+t+1}} = o(1)$ trivially implies $\|f(x_1+\dots+x_s, x_1+\dots+x_s+y_1+\dots+y_t)\|_{U^{s+t}} = o(1)$ (see the first step of proof of Corollary 3.3) and for $a(x) = x^{s+t-2}$, we have $\|f(x_1+\dots+x_s, x_1+\dots+x_s+y_1+\dots+y_t)\|_{U^{s+t}} = 1$. Finally, in the case $r=1$, $s,t \ge 1$ we can apply the same argument. Thus, the choice $r=0$, $s=1$, $t \ge 1$ is the only situation when we can obtain non–trivial upper bounds (under some analogs of assumption (35)) for $f(z_1+\dots+z_r+x_1+\dots+x_s, x_1+\dots+x_s+y_1+\dots+y_t)$.

Corollary 3.7 *Let K be a positive integer, $a : \mathbf{G} \to \mathbf{G}$ be a function, and $f(x,y) = e(x \cdot a(y))$. Suppose that*

$$|\{y \in \mathbf{G} \ : \ a(y) = j\}| \le K\,, \quad j \in \mathbf{G}\,. \tag{40}$$

Then

$$\|f(x,x+y)\|_{\mathcal{U}^2} \le 3^4 \max\left\{ 2^{30} \left(\frac{2^8}{\log(1/\varepsilon_*)} \right)^{2^{-19}}, 4 \left(\frac{2^8}{N} \right)^{2^{-16}} \right\} N^4\,,$$

where $\varepsilon_* = (K/N)^{1/4}$.

Proof. It is easy to see that

$$\|f(x,y)\|^4_{U^2} = \frac{1}{N^4} \sum_{x,x',y,y'} e((x-x')(a(y)-a(y'))) = \frac{1}{N^2} \sum_{j \in \mathbf{G}} |M_j|^2\,,$$

where $M_j = \{y \in \mathbf{G} \ : \ a(y) = j\}$. By assumption, we have $|M_j| \le K$ for all $j \in \mathbf{G}$. Besides $\sum_j |M_j| = N$. Hence $\|f(x,y)\|_{U^2} \le (K/N)^{1/4}$ and the statement follows from Corollary 3.3. This completes the proof.

Thus, if, say, $K = N^{1-\kappa}$, $N \to \infty$, $\kappa > 0$ then $\|f(x,x+y)\|_{U^2} \ll 1/(\log N)^c$, where $c = c(\kappa)$. Even in the situation when preimage of any (or almost any) point $j \in \mathbf{G}$ has the cardinality $o(N)$ we have a non–trivial bound for $\|f(x,x+y)\|_{U^2}$.

Note 3.8 It is easy to see that

$$\|e(xa(y_1 + \dots + y_t))\|_{\mathcal{U}^{t+1}} = N^2 \cdot \#\{y_1, \dots, y_t, y'_1, \dots, y'_t \ : \ \sum_{\omega \in \{0,1\}^t} a(y^\omega)(-1)^{|\omega|} = 0\}\,.$$

So, condition (35) can be interpreted as the requirement for $a(x)$ to be far from "polynomial of degree $t-1$" (or $(t-1)$–step nilsequence, more precisely, see e.g. [9, 24, 27]).

Finally, we apply Corollary 3.7 to a family of subsets of $\mathbf{G} \times \mathbf{G}$, $\mathbf{G} = \mathbb{Z}_N$. Let us recall some formulas from Fourier analysis. Let $f : \mathbf{G} \to \mathbb{C}$ be a function. By $\widehat{f}(\xi)$ denote the Fourier transformation of f

$$\widehat{f}(\xi) = \sum_{x \in \mathbf{G}} f(x) e(-\xi \cdot x)\,, \tag{41}$$

We have

$$\sum_{x \in \mathbf{G}} |f(x)|^2 = \frac{1}{N} \sum_{\xi \in \widehat{\mathbf{G}}} |\widehat{f}(\xi)|^2 \qquad \text{(Parseval identity)} \tag{42}$$

and

$$f(x) = \frac{1}{N} \sum_{\xi \in \widehat{\mathbf{G}}} \widehat{f}(\xi) e(\xi \cdot x) \qquad \text{(the inverse formula)}\ . \tag{43}$$

If

$$(f * g)(x) := \sum_{y \in \mathbf{G}} f(y) g(x-y)$$

then

$$\widehat{f * g} = \widehat{f}\widehat{g}\,. \tag{44}$$

Now we can formulate our corollary.

Corollary 3.9 *Let $\mathbf{G} = \mathbb{Z}_N$, N be a prime number. Let also $a : \mathbf{G} \to \mathbf{G}$ be a function, $a(y) \neq 0$ for all $y \in \mathbf{G}$. Suppose that P is an arithmetic progression, $|P| \geq 2^5 N^{3/4}$, $f(x,y) = P(x \cdot a(y)) - |P|/N$, and inequality (40) holds with $K \leq 2^{-18}|P|^4/N^3$. Then*

$$\|f(x, x+y)\|_{\mathcal{U}^2} \leq 3^4 \max\left\{ 2^{30} \left(\frac{2^8}{\log(1/\varepsilon_*)} \right)^{2^{-21}}, 4 \left(\frac{2^8}{N} \right)^{2^{-18}} \right\} N^4 , \tag{45}$$

where $\varepsilon_* = (4K/N)^{1/48}$.

Proof. To reduce some logarithms in our bounds we use a well–known trick. Without losing of generality, suppose that $P = \{0, 1, \dots, |P|-1\}$. Let $P_1 = \{0, 1, \dots, t-1\}$, $t = [|P|^{2/3} K^{1/12} N^{1/4} 2]$, and $W(x) = |P_1|^{-1}(P * P_1)(x)$. Clearly, $0 \leq W(x) \leq 1$, $\sum_x W(x) = |P|$ and $W(x) = P(x)$ for all but at most $2t$ points $x \in \mathbf{G}$. By Parseval, formula (44) and the Cauchy–Schwartz, we obtain

$$\sum_r |\widehat{W}(r)| \leq \frac{1}{|P_1|} \sqrt{N|P_1|} \sqrt{N|P|} = N|P|^{1/2} t^{-1/2} . \tag{46}$$

By the triangle inequality for $\mathcal{U}^2$—norm and (43), we get

$$\|f\|_{\mathcal{U}^2} \leq \|W - |P|/N\|_{\mathcal{U}^2} + 32 \cdot tN^3 =$$

$$= \frac{1}{N^4} \sum_{(r_1, \dots, r_4) \neq \vec{0}} \widehat{W}(r_1) \dots \widehat{W}(r_4) \sum_{x, x', y, y'} e(r_1 x a(y)) e(-r_2 x' a(y)) e(-r_3 x a(y')) e(r_4 x' a(y')) +$$

$$+32 \cdot tN^3 . \tag{47}$$

Clearly, $\|e(r \cdot x a(y))\|_{U^2} = \|e(x a(y))\|_{U^2}$, provided by $r \neq 0$. We have $\|e(x a(y))\|_{U^2} \leq (K/N)^{1/4}$ (see the proof of Corollary 3.7). Using the last inequality, Lemma 1.2, formulas (46), (47), we obtain

$$\|f\|_{\mathcal{U}^2} \leq \left(\frac{|P|}{t} \right)^2 \left(\frac{K}{N} \right)^{1/4} N^3 + 32 \cdot tN^3 \leq 2^8 \left(\frac{K}{N} \right)^{1/12} N^4 .$$

Using Corollary 3.3, we get (45). This concludes the proof.

Note 3.10 There is another alternative way to obtain the smallness of $\| \cdot \|_{U^2}$—norm of the function $f(x,y) = P(x \cdot a(y)) - |P|/N = P(x \cdot y) - |P|/N$ in the situation $\mathbf{G} = \mathbb{Z}_N$ and $a(y) = y$. It is easy to see, using multiplicative characters, that the $\|f\|_{U^2}$ is small by Pólya—Vinogradov inequality (see e.g. [1]).

References

[1] *Vinogradov I. M.* An introduction to the theory of numbers / Pergamon Press, London & New York, 1955.

[2] *Gowers W. T.* A new proof of Szemerédi's theorem for arithmetic progressions of length four // Geom. func. anal., v.8, 1998, 529–551.

[3] *Gowers W. T.* A new proof of Szemerédi's theorem // Geom. func. anal., v.11, 2001, 465–588.

[4] *Gowers W. T.* Quasirandomness, counting and regularity for 3–uniform hypergraphs // *Combin. Probab. Comput.* **15**:1–2 (2006) 143–184.

[5] *Gowers W. T.* Hypergraph Regularity and the multidimensional Szemerédi Theorem // arXiv:0710.3032v1 [math.CO] 16 Oct 2007.

[6] *P. Erdös, E. Szemerédi* On sums and products of integers // Studies in pure mathematics, 213–218, Birkhßauser, Basel, 1983.

[7] *J. Bourgain, N. Katz, T. Tao,* A sum-product estimate in finite fields and their applications // GAFA 14 (2004), N 1, 27–57.

[8] *Green B., Tao T.* The primes contain arbitrarily long arithmetic progressions // Annals of Math., submitted for publication, http:// www.arXiv:math.NT/0404188, 2004.

[9] *Green B., Tao T.* An inverse theorem for the Gowers $U^3(G)$ norm, with applications // http:// www.arXiv : math.NT/0503014 v1 1 Mar 2005.

[10] *Green B., Tao T.* The distribution of polynomials over finite fields, with aplications to the Gowers norms // arXiv:0711.3191v1 [math.CO] 20 Nov 2007.

[11] *Green B., Tao T.* The equivalence between inverse sumset theorems and inverse conjectures for the U^3—norm // arXiv:0906.3100v1 [math.NT] 17 Jun 2009.

[12] *Green B., Tao T.* The quantitative behaviour of polynomial orbits on nilmanifolds // arXiv:0709.3562v1 [math.NT] 22 Sep 2007.

[13] *Green B., Tao T.* Quadratic uniformity for the Möbius function // Annales de l'Institut Fourier (Grenoble) 58 (2008), no. 6, 1863–1935.

[14] *Green B.* Finite field model in additive combinatorics // Surveys in Combinatorics 2005, LMS Lecture Notes **329**, 1–29.

[15] *Tao T., Ziegler T.* The inverse conjecture for the Gowers norm over finite fields via the correspondence principle // preprint, available at arxiv.org/abs/0810.5527.

[16] *Samorodnitsky A., Trevisan L.* Gowers Uniformity, Influence of Variables, and PCPs // http:// www.arXiv : math.CO/0510264 v1 12 Oct 2005.

[17] *Lovett S., Meshulam R., Samorodnitsky A.* Inverse Conjecture for the Gowers norm is false // arXiv:0711.3388v1 [math.CO] 21 Nov 2007.

[18] *Shkredov I. D.* On a problem of Gowers // Izvestiya of Russian Academy of Scences, 70:2 (2006), 176–217.

[19] *Shkredov I. D.* On a Generalization of Szemerédi's Theorem // Proceedings London Math. Soc., 93:3 (2006), 723–760.

[20] *Shkredov I. D.* On a two–dimensional analogue of Szemerédi's theorem in Abelian groups // Izvestiya of Russian Academy of Scences, 73:5 (2009), 181–224.

[21] *Furstenberg H.* Recurrence in ergodic theory and combinatorial number theory / Princeton N.J., 1981.

[22] *Tao T.* A quantitative ergodic theory proof of Szemerédi's theorem // http:// www.arXiv : math.CO/0405251, 2004.

[23] *Tao T.* Norm convergence of multiple ergodic averages for commuting transformations // Ergodic Theory and Dynamical Systems, 28:657–688, 2008.

[24] *Host B., Kra B.* Nonconventional ergodic averages and nilmanifolds // Ann. Math., 161(1):397–488, 2005.

[25] *Host B., Kra B.* Convergence of Conze-Lesigne averages // Ergodic Theory Dynam. Systems, 21(2):493–509, 2001.

[26] *Host B., Kra B.* Convergence of polynomial ergodic averages // Israel J. Math., 149:1–19, 2005. Probability in mathematics.

[27] *Ziegler T.* A non-conventional ergodic theorem for a nilsystem // Ergodic Theory Dynam. Systems, 25(4):1357–1370, 2005.

[28] *Ziegler T.* Universal characteristic factors and Furstenberg averages // J. Amer. Math. Soc., 20(1):53–97 (electronic), 2007.

[29] *Austin T.* On the norm convergence of nonconventional ergodic averages // preprint.